[6] C. Y. Wu and K. N. Lai, "A new integrated voltage-controlled negative resistance device-LAMBDA MOSFET," *Solid-State Electron.*, vol. 22, pp. 779–781, Sep. 1979.

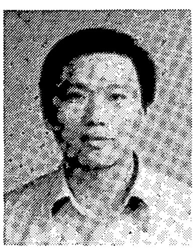
[7] C. Y. Wu and K. N. Lai, "Integrated Λ-type differential negative resistance MOSFET device," *IEEE J. Solid-State Circuits*, vol. SC-14, Dec. 1979.

[8] R. Genin and P. Brezel, "The generation of negative resistance by three-pole circuits," *Int. J. Electron.*, vol. 42, no. 6, pp. 589–600, 1977.

[9] L.O. Hill, D.O. Pederson, and R.S. Pepper, "Synthesis of electronic bistable circuits," *IEEE Trans. Circuit Theory*, vol. CT-10, pp. 25–35, Mar. 1963.

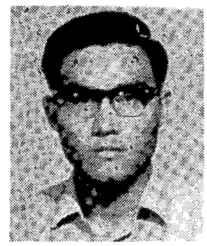
[10] H. Mizuno, G. Kano, H. Takagi, and I. Teramoto, "Theory of negative resistance of junction field-effect transistors," *IEEE J. Solid-State Circuits*, vol. SC-11, pp. 313–317, Apr. 1976.

[11] C. Y. Wu, K. N. Lai, and C. Y. Wu, "Generalized theory and realization of voltage-controlled negative resistance MOS device (LAMBDA MOSFET)," *Solid-State Electron.*, vol. 23, pp. 1–7, Jan. 1980.



Chung-Yu Wu (S'75–M'77) was born in Chiayi, Taiwan, Republic of China, on May 31, 1950. He received the B. S. degree, from the Department of Electrophysics, and the M. S. and Ph.D. degrees from Institute of Electronics., National Chiao-Tung University, Hsinchu, Taiwan, Republic of China, in 1972, 1976, and 1980, respectively.

During1975–1976, he studied ferroelectric films on silicon and their device applications. During 1976–1979, he had been engaged in the development of integrated differential negative resistance devices and their circuit applications, under the Grant support from the National Electronics Mass Plan (Semiconductor Devices and Integrated Circuit Technologies) of National Science Council. Since 1980 he has been an Associate Professor in Institute of Electronics. His current research interests have been in device modeling, IC technologies, and novel memory and analog IC design. He had received the 1976 Engineering Award for his M. S. thesis from the Electrical Engineering Society of Republic of China.



Ching-Yuan Wu (S'69–M'72) was born in Taiwan, on March 18, 1946. He received the B. S. degree from the Department of Electrical Engineering, National Taiwan University, Taiwan, in 1968, and the M. S. and Ph.D. degrees from State University of New York (SUNY), Stony Brook, in 1970 and 1972, respectively.

During the 1972–1973 academic year he was a Lecturer at the Electrical Sciences Department, SUNY, Stony Brook. In 1973 he was a Visiting Associate Professor at National Chiao-Tung University (NCTU), Taiwan, Republic of China. In 1976 he became a Professor in the Electronics Engineering Department, NCTU. During 1974–1980, he had been the Director of Engineering Laboratories and Semiconductor Research Center. He was the principal investigator of National Electronics Mass Plan–Semiconductor Devices and Integrated Circuit Technologies, during 1976–1979. Under a Grant support from National Science Council (NSC) and Ministry of Education (MOE), Republic of China, he designed and set up a complete research pilot line for integrated-circuit designs and fabrications. Currently, he is Professor and Director of the Institute of Electronics. He has been the Consultant of MOE in Sciences and Technology Education and the Research Consultant of the Electronics Research and Service Organization (ERSO), ITRI, Rep. of China. His research activities have been in integrated-circuit technologies, device modeling, and integrated-circuit designs. His present research interests focus on design and fabrication of high performance solar cells, silicon-on-sapphire, and CVD deposition for integrated- circuit applications, and new devices and new technologies.

Minimal Cascade Factorization of Real and Complex Rational Transfer Matrices

PAUL M. VAN DOOREN, MEMBER, IEEE, AND PATRICK DEWILDE, MEMBER, IEEE

Abstract—**A cascade factorization $R(\lambda)=R_1(\lambda)\cdot R_2(\lambda)\cdots R_k(\lambda)$ of an $n\times n$ nonsingular rational matrix $R(\lambda)$ is said to be minimal when the McMillan degrees of the factors add up to the McMillan degree of the original matrix. In this paper we give necessary and sufficient conditions for such a factorization to exist in terms of a state-space realization for $R(\lambda)$. Next, we focus on numerical and algorithmic aspects. We discuss the numerical conditioning of the problem and we give algorithms to compute degree one factorizations and real degree two factorizations. Finally, we discuss the special case where $R(\lambda)$ is a para J-unitary matrix.**

Manuscript received June 2, 1980; revised October 26, 1980. This work was supported in part by the National Science Foundation under Contracts MCS78-11985 and ENG78-10003, and by the U.S. Army Research Office under Contract DAAG29-79-C-0215.

P. Van Dooren is with Phillips Research Laboratory, Brussels, Belgium.
P. Dewilde is with the Department of Electrical Engineering, Delft University of Technology, Delft, The Netherlands.

I. Introduction

The problem of factorization of an $n\times n$ rational matrix $R(\lambda)$ into two factors $R(\lambda)=R_1(\lambda)\cdot R_2(\lambda)$ has recently received a lot of attention. In [10] sufficient conditions for the factorization to exist are given and it is shown by example that, in general, nontrivial factorizations (i.e., where the factors are not constant) may not exist. Independently, Van Dooren [25] and Bart *et al.* [2] gave necessary and sufficient conditions in terms of a state-space realization for $R(\lambda)$ [3]. The mathematical factorization theory is further elaborated in [2]. The work of Vandewalle [24] also deserves attention. There necessary and sufficient conditions are derived for a matrix $R_1(\lambda)$ to be a minimal factor of $R(\lambda)$ in terms of the parameters of the matrix itself.

0098-4094/81/0500-0390$00.75

In the past, cascade factorization has received a great deal of attention in circuit theory because of its theoretical and practical importance both in analysis and synthesis of networks [4], [9], [11]. While the solution of the problem of cascade factorization for special types of transfer functions has already been known for quite some time, it is only recently that one focused on the general case.

In this paper we give a complete solution of the cascade factorization problem using state-space techniques, thereby reviewing some recent results and streamlining the theory. During the elaboration of this paper, our attention has been drawn to the work of Sachnovic [17], which contains some of the results of this paper. Sachnovic, however, does not give proofs. We shall give simple proofs and next focus on numerical and algorithmic aspects of the problem.

Notations will be as follows. We reserve upper case for matrices and lower case for vectors and scalars. Script is used for operators and spaces. With $\mathcal{X}\cap\mathcal{Y}$ and $\mathcal{X}\oplus\mathcal{Y}$ we denote the intersection and the direct sum, respectively, of the spaces $\mathcal{X}$ and $\mathcal{Y}$. In the coordinate system used, $\mathcal{X}^{\perp}$ denotes the orthogonal complement of $\mathcal{X}$ and $\mathcal{X}\subset\mathcal{Y}$ means $\mathcal{X}$ is included in $\mathcal{Y}$. We use $\|\cdot\|_2$ and $\|\cdot\|_\infty$ for the 2-norm and ∞-norm of vectors and also for the corresponding induced matrix norms. $\mathbb{R}$ and $\mathbb{C}$ are the fields of real and complex numbers. The conjugate transpose of a matrix A is denoted by A^* and its largest and smallest singular values by $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$, respectively. When quantities are defined in an equality we use $\triangleq$.

In the next section we briefly review the geometrical results of [3] for minimal factorizations of the type $R(\lambda)=R_1(\lambda)\cdot R_2(\lambda)$. In Section III we show how to extend these results to general cascade factorizations involving an arbitrary number of factors. In Section IV we discuss the computational aspects of the problem and give numerical algorithms to perform the factorization for real and complex transfer functions (see also [25] [26]). Finally, in the last section, we connect our results with earlier work [6], [11] on para J-unitary transfer functions and show how they reduce to the classical theory due to Livsic [11].

II. Factorization Theory

Let $R(\lambda)$ be an $n\times n$ invertible rational matrix with coefficients in $\mathbb{C}$ or $\mathbb{R}$. The degree of the matrix is understood to be the Smith–McMillan degree and may be defined in terms of the poles of $R(\lambda)$ [12], [16]

Definition 1

The McMillan degree of a rational matrix is its number of poles, multiplicities counted. $\nabla\nabla\nabla$

There is some discussion as to how multiplicities must be counted. The conventional definition in dynamic system theory is the number of free modes of the system; this is also the definition adopted by McMillan. In this paper we shall assume that $R(\lambda)$ is given by a minimal state-space realization:

$$R(\lambda)=D+C(\lambda I_\delta-A)^{-1}B \tag{1}$$

whereby we tacitly assume that $R(\lambda)$ has no pole at infinity. Moreover, we shall take D to be nonsingular, so that $R(\lambda)$ does not have zeros at infinity as well. These assumptions do not restrict generality because a bilinear transformation on λ can always produce the desired form (1) and does not affect the factorizability. Under the assumptions made, a minimal realization for $R^{-1}(\lambda)$ is given by (see [16])

$$\begin{aligned}R^{-1}(\lambda)&=\hat{D}+\hat{C}(\lambda I_\delta-\hat{A})^{-1}\hat{B}\\ &\triangleq D^{-1}-D^{-1}C(\lambda I_\delta-A+BD^{-1}C)^{-1}BD^{-1}.\end{aligned} \tag{2}$$

The eigenvalues of A are the poles of $R(\lambda)$. The eigenvalues of $\hat{A}$ are the poles of $R^{-1}(\lambda)$ and thus the zeros of $R(\lambda)$. Poles and zeros are both δ in number, which is also the degree of $R(\lambda)$. The minimal state-space realization (1) will be denoted by the quadruple $\{A,B,C,D\}$. It is known (see, e.g., [16]) that all minimal realizations for $R(\lambda)$ are then given by $\{T^{-1}AT,T^{-1}B,CT,D\}$ with T an invertible transformation (also called a *state-space transformation*). We shall call realizations $\{A,B,C,D\}$ and $\{\hat{A},\hat{B},\hat{C},\hat{D}\}$ *corresponding* when they satisfy the relation (2). Under a state-space transformation T, corresponding realizations are transformed simultaneously as $\{T^{-1}AT,T^{-1}B,CT,D\}$ and $\{T^{-1}\hat{A}T,T^{-1}\hat{B},\hat{C}T,\hat{D}\}$. Hence we may consider coordinate-free operators $\mathcal{A}$ and $\hat{\mathcal{A}}$ acting in a state-space $\mathcal{X}$, which is either $\mathbb{C}^\delta$ or $\mathbb{R}^\delta$, and for which the poles and zeros of $R(\lambda)$ are the respective spectra. To any representation A and $\hat{A}$ of these operators in a given coordinate system, there correspond realizations $\{A,B,C,D\}$ and $\{\hat{A},\hat{B},\hat{C},\hat{D}\}$ for $R(\lambda)$ and $R^{-1}(\lambda)$ which are related as in (2). We call $\mathcal{A}$ and $\hat{\mathcal{A}}$ the *pole operator* and *zero operator*, respectively, of $R(\lambda)$.

Definition 2

The factorization

$$R(\lambda)=R_1(\lambda)\cdot R_2(\lambda) \tag{3}$$

is said to be minimal when the degrees δ_i of $R_i(\lambda)$ for $i=1,2$ add up to the degree δ of $R(\lambda)$. $\nabla\nabla\nabla$

It is well known (see, e.g., [16]) that for arbitrary factorizations of the type (3), one has

$$\delta\leqslant\delta_1+\delta_2. \tag{4}$$

The difference between the two sides is the number of pole-zero cancellations between R_1 and R_2. Minimality of a factorization is thus equivalent to the absence of cancellations between the two factors. The poles and zeros of both factors then constitute together the poles and zeros of the product $R(\lambda)$. This implies that both factors of the minimal factorization (3) have no poles nor zeros at infinity and that they must have minimal realizations:

$$R_1(\lambda)\triangleq D_1+C_1(\lambda I_{\delta_1}-A_{11})^{-1}B_1 \tag{5}$$

$$R_2(\lambda)\triangleq D_2+C_2(\lambda I_{\delta_2}-A_{22})^{-1}B_2 \tag{6}$$

with D_1 and D_2 constant and invertible. For notational convenience we use the so-called system matrices [16] to denote these realizations:

$$R_1(\lambda)\Leftrightarrow\left[\begin{array}{c|c}\lambda I_{\delta_1}-A_{11} & B_1\\ \hline -C_1 & D_1\end{array}\right] \quad (7a)$$

$$R_2(\lambda)\Leftrightarrow\left[\begin{array}{c|c}\lambda I_{\delta_2}-A_{22} & B_2\\ \hline -C_2 & D_2\end{array}\right]. \quad (7b)$$

The following is readily verified.

Lemma 1 (see, e.g., [7])

Let $R_i(\lambda)$ be realized as in (7). Then a realization for the product $R_1(\lambda)\cdot R_2(\lambda)$ is given by

$$R_1(\lambda)\cdot R_2(\lambda)\Leftrightarrow\left[\begin{array}{c:c|c}\lambda I_{\delta_1}-A_{11} & -B_1C_2 & B_1D_2\\ \hdashline 0 & \lambda I_{\delta_2}-A_{22} & B_2\\ \hline -C_1 & -D_1C_2 & D_1D_2\end{array}\right]. \quad (8)$$

▽▽▽

The following theorem is a corner stone for the rest of the paper (see also [3]):

Theorem 1

A rational matrix $R(\lambda)$ can be factorized minimally into factors of degree δ_1 and δ_2 iff it has a minimal realization in which the pole and zero matrices have the upper and lower block-triangular forms

$$A_f=\begin{bmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{bmatrix}\begin{matrix}\}\delta_1\\ \}\delta_2\end{matrix} \quad (9a)$$
$$\qquad\ \ \underbrace{\phantom{A_{11}}}_{\delta_1}\ \underbrace{\phantom{A_{22}}}_{\delta_2}$$

$$\hat{A}_f=\begin{bmatrix}\hat{A}_{11} & 0\\ \hat{A}_{21} & \hat{A}_{22}\end{bmatrix}\begin{matrix}\}\delta_1\\ \}\delta_2\end{matrix}. \quad (9b)$$
$$\qquad\ \ \underbrace{\phantom{A_{11}}}_{\delta_1}\ \underbrace{\phantom{A_{22}}}_{\delta_2}$$

Proof: Only if: Since there exists a minimal factorization, suppose that the factors $R_i(\lambda)$ are given by (7a) and (7b). Then (8) is a minimal realization for $R(\lambda)$, which is of the form (9a). It is readily verified that the zero matrix of (8) has the form (9b). Notice that $\hat{A}_{ii}=A_{ii}-B_iD_i^{-1}C_i$, so that A_{ii} and $\hat{A}_{ii}$ for $i=1,2$ are also the corresponding pole and zero matrices of (7a) and (7b).

If: Let the realization of $R(\lambda)$ corresponding to (9) be partitioned as

$$R(\lambda)\Leftrightarrow\left[\begin{array}{c:c|c}\lambda I_{\delta_1}-A_{11} & -A_{12} & B_{11}\\ \hdashline 0 & \lambda I_{\delta_2}-A_{22} & B_{22}\\ \hline -C_{11} & -C_{22} & D\end{array}\right]. \quad (10)$$

Then because $\hat{A}_{12}=0$ in (9b), we have that $A_{12}-B_{11}D^{-1}C_{22}=0$. The following matrix then has rank n and can be factorized as

$$\begin{bmatrix}-A_{12} & B_{11}\\ -C_{22} & D\end{bmatrix}\triangleq\underbrace{\begin{bmatrix}B_1\\ D_1\end{bmatrix}}_{n}\left[-C_2 \quad D_2\right]\}n. \quad (11)$$

Using Lemma 1 and (11) one can check that the factors

$$R_1(\lambda)\Leftrightarrow\left[\begin{array}{c|c}\lambda I_{\delta_1}-A_{11} & B_1\\ \hline -C_{11} & D_1\end{array}\right] \quad (12a)$$

$$R_2(\lambda)\Leftrightarrow\left[\begin{array}{c|c}\lambda I_{\delta_2}-A_{22} & B_{22}\\ \hline -C_2 & D_2\end{array}\right] \quad (12b)$$

yield the realization (10) for the product $R_1(\lambda)\cdot R_2(\lambda)$. This factorization is minimal because the realization (10) is minimal; hence $\delta=\delta_1+\delta_2$. ▽▽▽

Remark 1

The constant matrix factorization (11) is not unique. The matrix D_1 (or D_2) can be chosen arbitrarily but must be invertible since $D=D_1D_2$. The other three matrices then are defined by (11). This means that a constant (matrix) scaling factor can still be chosen arbitrarily between $R_1(\lambda)$ and $R_2(\lambda)$. For all possible factorizations (11), though, we still have that A_{ii} and $\hat{A}_{ii}$ are corresponding pole and zero matrices of the two factors (see only if part). ▽▽▽

In the sequel, we call a realization for $R(\lambda)$ *factorable* when A_f and $\hat{A}_f$ have the forms (9a) and (9b), since then realizations for the two factors are easily obtained by (11) and (12). The existence of such factorable realizations can geometrically be described by the invariant subspaces of the operators $\mathcal{A}$ and $\hat{\mathcal{A}}$ of $R(\lambda)$.

Theorem 2 (see also [3])

If $\mathcal{A}$ and $\hat{\mathcal{A}}$ are the pole and zero operators of $R(\lambda)$, then there exists a factorable realization for $R(\lambda)$ iff there exist independent subspaces $\mathcal{X}$ and $\mathcal{Y}$ of $\mathcal{H}$ such that:

$$\text{(i) } \mathcal{A}\mathcal{X}\subset\mathcal{X} \quad \text{(ii) } \hat{\mathcal{A}}\mathcal{Y}\subset\mathcal{Y} \quad \text{(iii) } \mathcal{X}\oplus\mathcal{Y}=\mathcal{H}. \quad (13)$$

Proof: If: Because of (iii) we can choose a coordinate system such that the columns of

$$\begin{bmatrix}I_{\delta_1}\\ 0\end{bmatrix}$$

form a basis for $\mathcal{X}$ and the columns of

$$\begin{bmatrix}0\\ I_{\delta_2}\end{bmatrix}$$

a basis for $\mathcal{Y}$. In this coordinate system $\mathcal{A}$ and $\hat{\mathcal{A}}$ are represented by matrices of the type (9a), (9b).

Only if: Let (9a), (9b) be the pole and zero operators of a factorable realization. In this coordinate system, the columns of

$$\begin{bmatrix}I_{\delta_1}\\ 0\end{bmatrix}$$

and

$$\begin{bmatrix}0\\ I_{\delta_2}\end{bmatrix}$$

are bases for invariant subspaces $\mathcal{X}$ and $\mathcal{Y}$ for $\mathcal{A}$ and $\hat{\mathcal{A}}$, respectively. Clearly condition (iii) is then also satisfied. ▽▽▽

Corollary 1

If $R(\lambda)$ has a realization $\{A, B, C, D\}$ and if the columns of X and Y are bases for subspaces $\mathcal{X}$ and $\mathcal{Y}$ satisfying Theorem 2, then $\{T^{-1}AT, T^{-1}B, CT, D\}$ is a factorable realization when $T \triangleq [X|Y]$, and only when T is of that type. $\nabla\nabla\nabla$

The factorization problem clearly is a combinatorial problem. For each pair of invariant subspaces $\mathfrak{X}$ and $\mathfrak{Y}$ of $\mathcal{Q}$ and $\hat{\mathcal{Q}}$ one can check for factorizability (condition (iii) in Theorem 2). The number of possible pairs is finite when $\mathcal{Q}$ and $\hat{\mathcal{Q}}$ do not have repeated eigenvalues but can be infinite otherwise (when $\mathcal{Q}$ or $\hat{\mathcal{Q}}$ has more than one Jordan block associated with a certain eigenvalue). The spaces $\mathfrak{X}$ and $\mathfrak{Y}$ also determine which will be the spectra of A_{ii} and $\hat{A}_{ii}$, after applying the transformation $T=[X|Y]$ of Corollary 1 that reduces the pole and zero matrices to the form (9). While $\mathfrak{X}$ thus chooses which poles will belong to R_1 and R_2, $\mathfrak{Y}$ performs the choice for the zeros. Hence for each pairs of subspaces $\mathfrak{X}$ and $\mathfrak{Y}$ there corresponds a certain distribution of poles and zeros over the two factors. Condition (iii) in Theorem 2 is a compatibility condition for the spaces $\mathfrak{X}$ and $\mathfrak{Y}$ saying whether or not there exists a factorization with such a distribution of poles and zeros. In the sequel we always work with one specific choice of $\mathfrak{X}$ and $\mathfrak{Y}$ and we discuss factorizability of $R(\lambda)$ in terms of that pair only.

When two bases X and Y of spaces $\mathfrak{X}$ and $\mathfrak{Y}$ with appropriate dimensions are thus put together in a matrix $T=[X|Y]$ then condition (iii) is satisfied if and only if T is invertible. In practice this condition is almost always satisfied. This has the rather annoying side-effect that any algorithm running on computer will not recognize unfactorable matrices since almost any small perturbation—caused by rounding errors in the computer—will make the matrix T invertible. However, closeness to an unfactorable matrix can still be detected. Indeed if the bases X and Y are very skew to each other we might as well say that the condition is "essentially" not satisfied. This is reflected in the numerical effort to be performed if one still wants to factorize such a matrix: the state-space transformation T to bring the original state-space model to a factorable form is then very badly conditioned. The conditioning of the transformation T is characterized by its condition number [8]:

$$\text{cond.}\, T \triangleq \|T\|_2 \cdot \|T^{-1}\|_2.$$

The "skewness" of two nontrivial subspaces $\mathfrak{X}$ and $\mathfrak{Y}$ is measured by their minimal angle $\theta_{\min}$, defined as

$$\begin{aligned}\cos\theta_{\min} &\triangleq \gamma_{\max}\\ &\triangleq \max\{|x^*y|\,|x\in\mathfrak{X}, y\in\mathfrak{Y}, \|x\|_2=\|y\|_2=1\},\\ &\qquad 0\leq\theta_{\min}\leq\pi/2. \qquad (14)\end{aligned}$$

The following theorem is proven in [3].

Theorem 3

(i) Any transformation $T=[X|Y]$ where the columns of X and Y are arbitrary bases for $\mathfrak{X}$ and $\mathfrak{Y}$ in a given coordinate system, satisfies

$$\text{cond.}\, T \geq \sqrt{\frac{1+\gamma_{\max}}{1-\gamma_{\max}}} = \frac{1+\cos\theta_{\min}}{\sin\theta_{\min}}. \qquad (15)$$

(ii) When X and Y are orthonormal bases in this coordinate system, this lower bound is met. $\nabla\nabla\nabla$

The optimal transformation to bring a state-space realization to factorable form is thus $T_0=[X_0|Y_0]$ where the columns of X_0 and Y_0 are orthonormal bases for $\mathfrak{X}$ and $\mathfrak{Y}$, respectively.

The point we wish to make is that an all-or-none factorizability condition is dangerous in practice, when inexact data and computations are involved, and that it is important to examine how close one is to satisfying or violating the condition. The above analysis shows that the minimal angle $\theta_{\min}$ gives a quantitative measure for this "closeness." When orthonormal bases X_0 and Y_0 are used for the spaces $\mathfrak{X}$ and $\mathfrak{Y}$ then $\|Y_0^*X_0\|_2=\gamma_{\max}=\cos\theta_{\min}$ [3]. When $\gamma_{\max}$ is not too close to 1, the transformation $T_0=[X_0|Y_0]$ is then used to construct a factorable realization, as recommended by Theorem 3. Since cond. T_0 tells us what accuracy can be expected after applying this transformation, we can determine from (15) which $\gamma_{\max}$ is still acceptable, depending on the requested accuracy (if, e.g., we want cond. $T_0\leq 100$ then $\gamma_{\max}$ should be smaller than 0.9998). For values of $\gamma_{\max}$ larger than this threshold we would call the rational matrix numerically unfactorable (as far as the considered pair $\mathfrak{X},\mathfrak{Y}$ is concerned).

It is to be noted that the choice of the initial state-space realization influences the condition of T. If no "natural" coordinate system is given, then it is wise to start out from a so-called balanced realization [14], for in that case the sensitivities of the parameters in the realization and the parameters in the transfer function are closely related.

III. Cascade Factorization

The problem here is to obtain a cascade factorization

$$R(\lambda)=R_1(\lambda)\cdot R_2(\lambda)\cdots R_k(\lambda) \qquad (16)$$

where all $R_i(\lambda)$ are of small degree, and the factorization is minimal in the following sense.

Definition 3

The cascade factorization (16) is termed minimal when the degrees δ_i of the factors $R_i(\lambda)$ add up to δ, the degree of $R(\lambda)$. $\nabla\nabla\nabla$

This means again that no pole-zero cancellations may occur in the cascaded product $R_1\cdot R_2\cdots R_k$, and that each of the factors must have a minimal realization

$$R_i(\lambda)\Leftrightarrow\left[\begin{array}{c|c}\lambda I_{\delta_i}-A_{ii} & B_i\\ \hline -C_i & D_i\end{array}\right] \qquad (17)$$

with D_i constant and invertible.

Lemma 2

Let the factors $R_i(\lambda)$ be realized by (17) for $i=1,\cdots,k$ and define

$$A_{ij}\triangleq B_iD_{i+1}\cdots D_{j-1}C_j, \qquad \text{for } i<j, \qquad (18a)$$

$$B_{jj}\triangleq B_jD_{j+1}\cdots D_k \qquad (18b)$$

$$C_{jj}\triangleq D_1\cdots D_{j-1}C_j, \qquad (18c)$$

$$D_{kk}\triangleq D_1\cdots D_k. \qquad (18d)$$

Then the product (16) is realized by $\{A, B, C, D\}$ with

$$A \triangleq \begin{bmatrix} A_{11} & \cdots & A_{1k} \\ & \ddots & \vdots \\ 0 & & A_{kk} \end{bmatrix} \tag{19a}$$

$$B \triangleq \begin{bmatrix} B_{11} \\ \vdots \\ B_{kk} \end{bmatrix} \begin{matrix} \}\delta_1 \\ \\ \}\delta_k \end{matrix} \tag{19b}$$

$$C \triangleq [\underbrace{C_{11}}_{\delta_1} \quad \cdots \quad \underbrace{C_{kk}}_{\delta_k}] \tag{19c}$$

$$D \triangleq [D_{kk}]. \tag{19d}$$

Proof: We prove the lemma by induction.

Lemma 1 proves the result for $k=2$. Assuming that the lemma is true for k, we prove now it also holds for $k' \triangleq k+1$. Since $R(\lambda)$ is realized by $\{A, B, C, D\}$ as in (19) and $R_{k'}(\lambda)$ by $\{A_{k'k'}, B_{k'}, C_{k'}, D_{k'}\}$, a realization $\{A', B', C', D'\}$ for their product $R'(\lambda)=R(\lambda)\cdot R_{k'}(\lambda)$ can be obtained by Lemma 1 as follows:

$$A' \triangleq \left[\begin{array}{c|c} A & BC_{k'} \\ \hline 0 & A_{k'k'} \end{array}\right]$$

$$B' \triangleq \left[\begin{array}{c} BD_{k'} \\ \hline B_{k'} \end{array}\right] \begin{matrix} \}\delta \triangleq \sum_{i=1}^{k} \delta_i \\ \}\delta_{k'} \end{matrix}$$

$$C' \triangleq [\underbrace{C}_{\delta} \mid \underbrace{DC_{k'}}_{\delta_{k'}}]$$

$$D' \triangleq [DD_{k'}].$$

This verifies (18) and (19) for k updated to k'. $\nabla\nabla\nabla$

This leads to the following generalization of Theorem 1.

Theorem 4

A rational matrix has a minimal cascade factorization (16) into factors of degrees δ_i, $i=1,\cdots,k$ iff it has a minimal realization in which the pole and zero matrices have the upper and lower block-triangular forms:

$$A_c \triangleq \begin{array}{c} \begin{bmatrix} A_{11} & X & \cdots & X \\ & A_{22} & \cdots & X \\ 0 & & \ddots & \vdots \\ & & & A_{kk} \end{bmatrix} \begin{matrix} \}\delta_1 \\ \}\delta_2 \\ \vdots \\ \}\delta_k \end{matrix} \\ \begin{matrix} \delta_1 & \delta_2 & \cdots & \delta_k \end{matrix} \end{array} \tag{20a}$$

$$\hat{A}_c \triangleq \begin{array}{c} \begin{bmatrix} \hat{A}_{11} & & 0 & \\ X & \hat{A}_{22} & & \\ \vdots & \vdots & \ddots & \\ X & X & \cdots & \hat{A}_{kk} \end{bmatrix} \begin{matrix} \}\delta_1 \\ \}\delta_2 \\ \vdots \\ \}\delta_k \end{matrix} \\ \begin{matrix} \delta_1 & \delta_2 & \cdots & \delta_k \end{matrix} \end{array}. \tag{20b}$$

Proof: Only if: Let $R(\lambda)$ have a minimal factorization (16), then (19) gives a minimal realization for $R(\lambda)$. It is easy to see that (18) yields the identities

$$A_{ij} - B_{ii}D_{kk}^{-1}C_{jj} = 0, \qquad \text{for } i<j \tag{21}$$

which are exactly the upper diagonal blocks $\hat{A}_{ij}$ of the zero matrix $\hat{A}=A-BD^{-1}C$ of (19).

If: Let $\{A, B, C, D\}$ be a realization such that A and $\hat{A}=A-BD^{-1}C$ have the form (20). Then we can partition this realization as in (19). One can always construct factors $R_i(\lambda)$ with realizations $\{A_{ii}, B_i, C_i, D_i\}$ satisfying (18), (19) as follows. Choose k invertible factors D_i that satisfy (18d) and solve for B_i and C_i (for $j=1,\cdots,k$) from (18b) and (18c); finally (18a) is also satisfied because of (21), indicating that $\hat{A}$ is lower triangular. This factorization is minimal because the realization $\{A, B, C, D\}$ is minimal and thus

$$\delta = \sum_{i=1}^{k} \delta_i. \qquad \nabla\nabla\nabla$$

Remark 2

Just as in the case of single factorizations we can choose $(k-1)$ of the matrices D_i arbitrarily (if nonsingular) in (18d). This is equivalent to saying that $(k-1)$ constant scaling factors can arbitrarily be chosen between the pairs $R_i \cdot R_{i+1}$ (for $i=1,\cdots,k-1$). Yet, for any such choice we still have from (18) that

$$\hat{A}_{ii} = A_{ii} - B_i D_i^{-1} C_i$$

which means that the diagonal blocks A_{ii} and $\hat{A}_{ii}$ of (20) are the pole and zero matrices of the factors $\{A_{ii}, B_i, C_i, D_i\}$. $\nabla\nabla\nabla$

To avoid too much computations one should choose, e.g., $D_1=D$ and the other $D_i=I$ which then gives the following construction for the factors of $R(\lambda)$ realized by (19):

$$\begin{cases} R_1(\lambda) \text{ is realized by } \{A_{11}, B_{11}, C_{11}, D\} \\ R_i(\lambda) \text{ is realized by } \{A_{ii}, B_{ii}, D^{-1}C_{ii}, I\}, \text{ for } i=2,\cdots,k. \end{cases}$$

Since the factors $R_i(\lambda)$ can easily be obtained from the realization of $R(\lambda)$ when it has the property (20), we will call such a realization *cascaded*. Necessary and sufficient conditions for the existence of such cascaded realizations are given by:

Theorem 5

If $\mathcal{A}$ and $\hat{\mathcal{A}}$ are the pole and zero operators of $R(\lambda)$, then there exists a cascaded realization (20) for $R(\lambda)$ iff there exist $(k-1)$ pairs of independent subspaces $\mathcal{X}_i, \mathcal{Y}_i$, nested as follows:

$$\{0\} \subset \mathcal{X}_1 \subset \cdots \subset \mathcal{X}_{k-1} \subset \mathcal{H} \tag{22a}$$

$$\mathcal{H} \supset \mathcal{Y}_1 \supset \cdots \supset \mathcal{Y}_{k-1} \supset \{0\} \tag{22b}$$

and such that the following conditions are satisfied for each pair:

$$\text{(i)} \quad \mathcal{A}\mathcal{X}_i \subset \mathcal{X}_i \qquad \text{(ii)} \quad \hat{\mathcal{A}}\mathcal{Y}_i \subset \mathcal{Y}_i \qquad \text{(iii)} \quad \mathcal{X}_i \oplus \mathcal{Y}_i = \mathcal{H}. \tag{23}$$

Proof: If: Let us define $\mathcal{X}_k \triangleq \mathcal{Y}_0 \triangleq \mathcal{K}$, $\mathcal{X}_0 \triangleq \mathcal{Y}_k \triangleq \{0\}$, and $\mathbb{S}_i \triangleq \mathcal{X}_i \cap \mathcal{Y}_{i-1}$ for $i=1,\cdots,k$. It follows then from (22) and (23iii) that

$$\begin{cases} \mathcal{X}_i = \mathcal{X}_{i-1} \oplus \mathbb{S}_i \\ \mathcal{Y}_i = \mathbb{S}_{i+1} \oplus \mathcal{Y}_{i+1}, & \text{for } i=1,\cdots,k-1 \end{cases}$$

and hence, by induction, that

$$\begin{cases} \mathcal{X}_i = \mathbb{S}_1 \oplus \cdots \oplus \mathbb{S}_i \\ \mathcal{Y}_i = \mathbb{S}_{i+1} \oplus \cdots \oplus \mathbb{S}_k, & \text{for } i=1,\cdots,k-1. \end{cases} \tag{24}$$

Let ξ_i, η_i, and δ_i be the dimensions of $\mathcal{X}_i$, $\mathcal{Y}_i$, and $\mathbb{S}_i$, respectively. It follows from (23iii) that

$$\mathbb{S}_1 \oplus \mathbb{S}_2 \oplus \cdots \oplus \mathbb{S}_k = \mathcal{K}.$$

We thus can choose a coordinate system in which $\mathbb{S}_i$ is spanned by the columns of

$$\begin{bmatrix} 0 \\ I_{\delta_i} \\ 0 \end{bmatrix} \begin{matrix} \}\xi_{i-1} \\ \\ \}\eta_i \end{matrix}, \qquad \text{for } i=1,\cdots,k \tag{25}$$

whence $\mathcal{X}_i$ and $\mathcal{Y}_i$ are spanned by the columns of

$$\begin{bmatrix} I_{\xi_i} \\ 0 \end{bmatrix}$$

and

$$\begin{bmatrix} 0 \\ I_{\eta_i} \end{bmatrix}$$

respectively.

In this coordinate system the pole and zero matrices A and $\hat{A}$ have the required form (20) because of conditions (23i) and (23ii).

Only if: In the coordinate system of (20) the spaces $\mathcal{X}_i$ and $\mathcal{Y}_i$ defined as in (24),(25) clearly satisfy all conditions. ∇∇∇

Corollary 2

If $R(\lambda)$ has a realization $\{A,B,C,D\}$ and if in this coordinate system the columns of S_i are bases for the subspaces $\mathbb{S}_i$ of Theorem 5, then $\{T^{-1}AT, T^{-1}B, CT, D\}$ is a cascaded realization when $T=[S_1|\cdots|S_k]$, and only when T is of that type.

Proof: Under this transformation new bases for $\mathbb{S}_i$ are given by (25) as required in Theorem 5 (if and only if). ∇∇∇

Note that the obtained factors $R_i(\lambda)$ will be nontrivial ($\delta_i \neq 0$) if and only if the inclusions in (22) are strict. This is always assumed in the sequel. The combinatorial aspect of the problem is obvious again. Each pair of chains (22) of subspaces $\mathcal{X}_i, \mathcal{Y}_i$, invariant under $\mathcal{Q}$ and $\hat{\mathcal{Q}}$, respectively—i.e., satisfying (23i) and (23ii)—partitions the poles and zeros of $R(\lambda)$ over the factors $R_i(\lambda)$. Condition (23iii), finally, says whether or not a cascade factorization can be obtained using the subspaces $\{\mathcal{X}_i, \mathcal{Y}_i | i=1,\cdots,k\}$. As in Section II we replace this last condition by a quantitative measure which is a more realistic criterion for numerical practice. This is done as follows. For a given chain of invariant subspaces (22),(23i),(23ii) we can construct the spaces $\mathbb{S}_i$ of Theorem 5. Condition (23iii) is then satisfied if and only if these spaces are linearly independent and span the whole space $\mathcal{K}$. When, in a given metric, some of the $\mathbb{S}_i$ are too skew to each other, we may say that condition (23iii) is "almost" not satisfied. This is again reflected in the condition number of the transformation $T=[S_1|\cdots|S_k]$ required by Corollary 2. Although one would expect from Theorem 3 that cond. T is minimal when the bases S_i are chosen orthonormal, this does not hold anymore when $k>2$ (see [22]). Yet, a straightforward construction of a transformation of the type $T=[S_1|\cdots|S_k]$ with minimal condition number is not known [22]. The following theorem shows that using orthonormal bases S_i gives a relatively low condition number.

Theorem 6

Let $T_0=[S_1^0|\cdots|S_k^0]$ be an invertible matrix where the columns of S_i^0 are an orthonormal basis for $\mathbb{S}_i$. Let $\theta_{\min}$ be the minimal angle between any two of the subspaces $\mathbb{S}_i$, and let $\mathcal{D}_k$ be the class of nonsingular block diagonal matrices partitioned conforming to T_0. Then

$$\text{cond.}\, T_0 \leq \sqrt{1+(k-1)\cos\theta_{\min}} \cdot \inf_{\mathcal{D}\in\mathcal{D}_k} \text{cond.}\, T_0 D.$$

Proof: Let $D \triangleq \operatorname{diag}\{D_1,\cdots,D_k\}$ be an arbitrary matrix in $\mathcal{D}_k$. Then the following is readily checked:

$$\|T_0 D\|_2 \geq \max_j \|S_j^0 D_j\|_2 = \max_j \|D_j\|_2 = \|D\|_2$$

and, therefore,

$$\begin{aligned} \text{cond.}(T_0 D) &= \|T_0 D\|_2 \|D^{-1} T_0^{-1}\|_2 \\ &\geq \|D\|_2 \|D^{-1} T_0^{-1}\|_2 \geq \|T_0^{-1}\|_2. \end{aligned} \tag{26a}$$

Next, let x be any vector of unit norm partitioned conforming to T_0:

$$x' = (x_1', \cdots, x_k').$$

Then

$$\begin{aligned} \|T_0 x\|_2^2 &= \left\| \sum_{j=1}^k S_j^0 x_j \right\|_2^2 = \left| \sum_{i=1}^k \sum_{j=1}^k x_i^* S_i^{0*} S_j^0 x_j \right| \\ &\leq \sum_{i=1}^k |x_i^* x_i| + \sum_{i=1}^k \sum_{\substack{j=1 \\ j\neq i}}^k \|x_i\|_2 \|x_j\|_2 \cos\theta_{\min} \\ &= 1 + \left[\sum_{i=1}^k \sum_{\substack{j=1 \\ j\neq i}}^k \|x_i\|_2 \|x_j\|_2 \right] \cos\theta_{\min} \\ &= 1 + \left\{ \left[\sum_{i=1}^k \sum_{j=1}^k \|x_i\|_2 \|x_j\|_2 \right] - 1 \right\} \cos\theta_{\min} \\ &= 1 + \left\{ \left[\sum_{i=1}^k \|x_i\|_2 \right]^2 - 1 \right\} \cos\theta_{\min}. \end{aligned}$$

Notice now that the sum in the last expression reaches its maximum when all $\|x_i\|_2$ are equal to $1/\sqrt{k}$, so that

$$\|T_0\|_2^2 = \max_x \|T_0 x\|_2^2 \leqslant 1+(k-1)\cos\theta_{\min}$$

and, therefore,

$$\text{cond.}(T_0) = \|T_0\|_2 \|T_0^{-1}\|_2 \leqslant \sqrt{1+(k-1)\cos\theta_{\min}}\, \|T_0^{-1}\|_2. \tag{26b}$$

Since (26a) holds for any D we obtain the desired result from (26a),(26b). ▽▽▽

Since the transformations $\{T_0 D \mid D \in \mathcal{D}_k\}$ are exactly those derived by Corollary 2, this theorem says that T_0 has almost minimal condition number among all suitable transformations. A measure for factorizability is thus given by cond. T_0. An ideal situation would be that the $\mathcal{S}_i$ are orthogonal in the given metric: then T_0 is orthonormal and cond. $T_0 = 1$ is minimal. When, on the other hand, cond. T_0 becomes very large and above some treshold, one should rather not perform the transformation T_0 because of its inherent numerical difficulty. Condition (23iii) is then "essentially" not satisfied.

IV. Computational Aspects

In this section we discuss some practical aspects of the theory developed in previous sections. For reasons that will become clear, we first focus on first degree factorizations:

$$R(\lambda) = R_1(\lambda)\cdot R_2(\lambda) \cdots R_\delta(\lambda) \tag{27}$$

where thus each factor $R_i(\lambda)$ has only one pole and one zero. We start with a given minimal realization $\{A, B, C, D\}$ for $R(\lambda)$. For the construction of a transformation T that reduces this realization to cascaded form, we need to compute invariant subspaces of the matrices A and $\hat{A} = A - BD^{-1}C$. A common way to proceed is to decompose A and $\hat{A}$ into upper and lower Schur form, respectively,

$$U^* A U \triangleq A_s \triangleq \begin{bmatrix} a_{11} & & x \\ & \ddots & \\ 0 & & a_{\delta\delta} \end{bmatrix} \tag{28a}$$

$$V^* \hat{A} V \triangleq \hat{A}_s \triangleq \begin{bmatrix} \hat{a}_{11} & & 0 \\ & \ddots & \\ x & & \hat{a}_{\delta\delta} \end{bmatrix} \tag{28b}$$

where U and V are unitary. Notice that $\{a_{ii}\}$ are the poles of $R(\lambda)$ and $\{\hat{a}_{ii}\}$ its zeros. Let us choose $\mathcal{X}_i$ as the space spanned by the first i columns of U and $\mathcal{Y}_i$ as the space spanned by the last $(\delta - i)$ columns of V, this for $i = 1,\cdots,\delta-1$. Then by (28) $\mathcal{X}_i$ and $\mathcal{Y}_i$ are invariant subspaces of A and $\hat{A}$, respectively, and they satisfy conditions (23i) and (23ii) by construction. The last condition (23iii) can then be expressed in terms of the leading principal minors of the matrix

$$Q \triangleq V^* U. \tag{29}$$

Indeed, let U and V be partitioned as

$$U \triangleq [\underbrace{U_1^i}_{i} \mid \underbrace{U_2^i}_{\delta-i}]\,; \quad V \triangleq [\underbrace{V_1^i}_{i} \mid \underbrace{V_2^i}_{\delta-i}]. \tag{30}$$

Then the columns of U_1^i, U_2^i, V_1^i, V_2^i are orthonormal bases for the spaces $\mathcal{X}_i$, $\mathcal{X}_i^\perp$, $\mathcal{Y}_i^\perp$, and $\mathcal{Y}_i$, respectively. We then have that (see, e.g., [5])

$$\sigma_{\max}\{V_1^{i*} U_2^i\} = \sigma_{\max}\{V_2^{i*} U_1^i\} = \cos\theta_{\min}^{(i)}$$
$$\sigma_{\min}\{V_1^{i*} U_1^i\} = \sigma_{\min}\{V_2^{i*} U_2^i\} = \sin\theta_{\min}^{(i)} \tag{31}$$

wherein $\theta_{\min}^{(i)}$ is the minimal angle between $\mathcal{X}_i$ and $\mathcal{Y}_i$. Hence, if (23iii) is satisfied, then $\theta_{\min}^{(i)} > 0$ and the leading principal minors $Q_{ii} = V_1^{i*} U_1^i$ of Q are nonsingular. This implies that there exists a LU factorization (see, e.g., [28], p. 204):

$$Q \triangleq Q_l \cdot Q_u \tag{32}$$

where Q_l and Q_u are lower and upper triangular, respectively. Moreover, we can always normalize the columns of Q_l and, e.g., require that its diagonal elements are positive real, which makes the decomposition unique. Then, after the state-space transformation

$$T \triangleq V \cdot Q_l = U \cdot Q_u^{-1} \tag{33}$$

it is easy to see that the new pole and zero matrices are upper and lower triangular, respectively,

$$A_c \triangleq T^{-1} A T = Q_u U^* A U Q_u^{-1} = Q_u A_s Q_u^{-1} = [\,\overline{\diagdown}\,] \tag{34a}$$

$$\hat{A}_c \triangleq T^{-1} \hat{A} T = Q_l^{-1} V^* \hat{A} V Q_l = Q_l^{-1} \hat{A}_s Q_l = [\,\underline{\diagdown}\,]. \tag{34b}$$

Note that the diagonal elements of (34a), (34b) are those of (28a), (28b) and that (34) indicates that $\{T^{-1}AT, T^{-1}B, CT, D\}$ is in cascaded form. When also reversing the above reasoning we finally obtain:

Theorem 7

A minimal cascade factorization (27) with given ordering of poles and zeros exists iff there exists an LU decomposition for the matrix $Q = V^* U$ where U and V are the unitary matrices which transform A and $\hat{A}$ to their Schur form (28) with the prescribed ordering of poles and zeros. ▽▽▽

Moreover, according to (33) and the normalization of Q_l, we have that the columns of T are normalized, whence T satisfies the conditions of Theorem 6. It is superfluous to prove that the columns of T indeed span the $\mathcal{S}_i$ spaces of Theorem 5 since $\{T^{-1}AT, T^{-1}B, CT, D\}$ is in cascaded form. The decompositions (28) and (32) thus yield a simple construction for a near optimal transformation T to reduce a realization to cascaded form. The condition number of T yet depends on the chosen Schur decompositions and hence on the ordering of the poles and zeros over the factors $R_i(\lambda)$. This is illustrated by the following simple example (here we have only two factors and an optimal T can be constructed for each pole–zero ordering):

Example 1

The transfer function

$$R(\lambda) = \begin{bmatrix} \dfrac{\lambda}{\lambda - 1} & 2 \\ 0 & \dfrac{\lambda - 1}{\lambda} \end{bmatrix}$$

has the minimal realization

$$R(\lambda) \Leftrightarrow \left[\begin{array}{cc|cc} \lambda & 0 & 0 & 1 \\ 0 & \lambda-1 & 1 & 0 \\ \hline 0 & -1 & 1 & 2 \\ 1 & 0 & 0 & 1 \end{array}\right].$$

The pole and zero operators are

$$A=\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \hat{A}=\begin{bmatrix} 1 & 0 \\ -2 & 0 \end{bmatrix}.$$

Their Schur decompositions are (with $c \triangleq 2/\sqrt{5}$; $s \triangleq 1/\sqrt{5}$)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \hat{A} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 0 \end{bmatrix}$$

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \hat{A} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -2 & 1 \end{bmatrix}.$$

For different orderings of poles $\{a_{ii}\}$ and zeros $\{\hat{a}_{ii}\}$ we give below the matrix Q to be factored as in (32) and the state-space transformation T derived therefrom

a_{11}	a_{22}	$\hat{a}_{11}$	$\hat{a}_{22}$	$Q=V^*U$	T	cond. T
0	1	1	0	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	1
0	1	0	1	$\begin{bmatrix} c & s \\ -s & c \end{bmatrix}$	$\begin{bmatrix} 1 & -s \\ 0 & c \end{bmatrix}$	$\frac{1+\sqrt{5}}{2}$
1	0	0	1	$\begin{bmatrix} s & c \\ c & -s \end{bmatrix}$	$\begin{bmatrix} 0 & -s \\ 1 & c \end{bmatrix}$	$2+\sqrt{5}$
1	0	1	0	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	/	(∞).

For the last combination, T does not exist. Constructing T as in Corollary 2 would indeed yield a singular T (i.e., cond. $T=\infty$) which is unacceptable. $\nabla\nabla\nabla$

Algorithms to perform the decompositions described above, can be found in the literature [18], [29]. The Schur decompositions (28a) and (28b) are obtained by the QR and QL algorithms, respectively, which are known to be numerically stable [28]. In order to obtain all possible orderings of poles and zeros in these decompositions, one can make use of an efficient and stable updating of these decompositions [21]. The LU factorization (32) is obtained by Gaussian elimination, but without pivoting. This is the only step that could be numerically unstable since no pivoting is allowed, but instability can only occur when cond. T is large, hence when the factorization becomes "essentially" impossible. As mentioned above, only transformations T with "reasonable" condition number should be allowed, say with cond. T smaller than some given K (e.g., cond. $T \leqslant 100$). Since

$$\text{cond.}\, T = \text{cond.}\, Q_l = \text{cond.}\, Q_u = \|Q_l\|_2 \|Q_u\|_2 \leqslant K \tag{35}$$

the factorization $Q=Q_l \cdot Q_u$ is also well conditioned in such a case and numerical stability can be ensured for its computation. The pivots q_{ii}^l in Q_l can then be bounded from below (ensuring the numerical stability):

$$|q_{ii}^l| \geqslant \sigma_{\min}\{Q_l\} \geqslant \|Q_l\|_2/K \geqslant 1/K$$

since they are the eigenvalues of Q_l. Note also that cond. $T \leqslant K$ is satisfied when (see, e.g., [19] p. 183)

$$\|Q_l\|_\infty \|Q_u\|_\infty \leqslant K/\delta$$

which is easily checked while computing the factorization of Q. In some cases it is preferable to compute a cascade factorization

$$R(\lambda)=R_1(\lambda)\cdot R_2(\lambda) \cdots R_k(\lambda) \tag{36}$$

where the separate factors $R_i(\lambda)$ have degrees larger than one; e.g., when $R(\lambda)$ is real one would prefer the factors to be real also and these will have degree at least 2 in general [23]. To compute the chains of invariant subspaces (22) we use the "block" Schur decompositions:

$$U^*AU \triangleq A_s \triangleq \begin{bmatrix} B_{11} & & X \\ & \ddots & \\ 0 & & B_{kk} \end{bmatrix} \tag{37a}$$

$$V^*\hat{A}V \triangleq \hat{A}_s \triangleq \begin{bmatrix} \hat{B}_{11} & & 0 \\ & \ddots & \\ X & & \hat{B}_{kk} \end{bmatrix} \tag{37b}$$

where U and V are unitary and B_{ii}, $\hat{B}_{ii}$ have dimensions $\delta_i \times \delta_i$. Here we choose $\mathcal{X}_i$ and $\mathcal{Y}_i$ to be the spaces spanned by the first i "block" columns of U and by the last $(k-i)$ "block" columns of V, respectively. By similar arguments to (28)–(34) we have that this choice of spaces $\mathcal{X}_i, \mathcal{Y}_i$ $i=1,\cdots,k-1$ satisfies the factorizability condition (23iii) if and only if the leading principal minors of $Q=V^*U$ that correspond to the block partitioning of (37), are nonsingular. We then decompose Q, analogously to (32), in a product of respectively lower and upper block triangular matrices partitioned conformably with (37):

$$Q \triangleq Q_l \cdot Q_u. \tag{38}$$

By analogy to (33)–(34), the state-space transformation

$$T \triangleq V \cdot Q_l = U \cdot Q_u^{-1} \tag{39}$$

is then shown to reduce A and $\hat{A}$ to A_c and $\hat{A}_c$, respectively, which have the required block triangular forms of Theorem 4. Note that the corresponding diagonal blocks of A and A_c (resp. $\hat{A}$ and $\hat{A}_c$) are similar and thus have the same spectrum. Similarly to Theorem 7, this leads to:

Theorem 8

A minimal cascade factorization (36) with specified block ordering of poles and zeros exists iff there exists a "block" LU decomposition of $Q=V^*U$ where U and V are the unitary matrices which transform A and $\hat{A}$ to their block Schur form (37) with the prescribed block ordering of poles and zeros. $\nabla\nabla\nabla$

Moreover, if the block columns of Q_l are normalized in the decomposition (38), this also holds for T in (39), whence T satisfies the conditions of Theorem 6.

In practice the decompositions (37) are computed using the QR and QL algorithms [29], yielding 1×1 diagonal blocks when working in complex arithmetic and 1×1 and 2×2 diagonal blocks when working in real arithmetic. If larger blocks are requested, this is obtained by appropriate block partitioning of A_s and $\hat{A}_s$. The distribution of eigenvalues over the diagonal blocks B_{ii} and $\hat{B}_{ii}$ can be obtained by reordering the QR and QL decompositions [21]. Numerically stable software is available for these two steps [21], [18]. The block LU decomposition (38) is performed by Gaussian elimination but with pivoting only allowed within the $\delta_i \times \delta_i$ blocks. In practice one would aim for a cascade factorization with as small as possible degrees δ_i for each factor (i.e., $\delta_i = 1$ in complex arithmetic and $\delta_i = 1$ or 2 in real arithmetic if complex conjugate eigenvalues occur). If within these constraints the pivots of the LU decomposition are too small—i.e., cond. Q_l will be too large—then larger $\delta_i \times \delta_i$ blocks are considered so as to allow more flexibility in the pivot choice.

Example 2

Let

$$R(\lambda) \Leftrightarrow \left[\begin{array}{cccc|cc} \lambda-3 & 0 & 2 & -1 & 1 & 0 \\ 0 & \lambda-2 & 2 & 1 & 1 & 1 \\ 0 & 0 & \lambda & 1 & 1 & 0 \\ 0 & 0 & -1 & \lambda & 0 & 1 \\ \hline -1 & 0 & 2 & -1 & 1 & 2 \\ 0 & -1 & 1 & 0 & 0 & -1 \end{array}\right]. \quad (40)$$

The pole and zero operators of $R(\lambda)$ in this coordinate system are

$$A = \left[\begin{array}{c:c:cc} 3 & 0 & -2 & 1 \\ \hdashline 0 & 2 & -2 & -1 \\ \hdashline 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

$$\hat{A} = \begin{bmatrix} 2 & -2 & 2 & 0 \\ -1 & 1 & 1 & -2 \\ -1 & -2 & 4 & -2 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

A is already in the so-called "real" Schur form and has eigenvalues $\{3, 2, i, -i\}$. Hence we can take $A_s = A$ and $U = I$. A possible Schur decomposition for $\hat{A}$, with eigenvalues ordered in decreasing norm, is (with $d \triangleq \sqrt{2}$):

$$\hat{A}_s \triangleq V^* \hat{A} V = \left[\begin{array}{c:c:cc} 3 & & & \\ \hdashline -2d & 2 & & \\ \hdashline -3 & -d & 2 & -2d \\ 1/d & 0 & 1/d & 0 \end{array}\right],$$

$$\text{with } V \triangleq \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/d & 0 & 1/d & 0 \\ -1/d & 0 & 1/d & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and with eigenvalues $\{3, 2, 1+i, 1-i\}$. The matrices A_s and $\hat{A}_s$ are partitioned conformably ($\delta_1 = 1$, $\delta_2 = 1$, $\delta_3 = 2$) and will lead to a cascade factorization with corresponding degrees δ_i if and only if the matrix $Q = V^*$ has a conformable block LU decomposition. Unfortunately, the element q_{11}, supposed to be the first pivot in this decomposition, is zero. Using a larger block partitioning ($\delta_1 = 2$, $\delta_2 = 2$) we obtain

$$Q = Q_l \cdot Q_u \triangleq \left[\begin{array}{cc:cc} 0 & 1/d & & \\ 1 & 0 & & \\ \hdashline 0 & 1/d & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \cdot \left[\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ \hdashline & & d & 0 \\ & & 0 & 1 \end{array}\right].$$

Here Q_l has orthonormal block columns and cond. $Q_l = \sqrt{2} + 1$. After applying the transformation $T \triangleq V \cdot Q_l$ to the state-space system (40), we obtain the factorable form:

$$R(\lambda) \Leftrightarrow \left[\begin{array}{cc:cc|cc} \lambda-3 & 0 & d & -1 & 1 & 0 \\ 0 & \lambda-2 & 0 & 0 & 0 & 1 \\ \hdashline 0 & 0 & \lambda & d & d & 0 \\ 0 & 0 & -1/d & \lambda & 0 & 1 \\ \hline -1 & 0 & d & -1 & 1 & 2 \\ 0 & -1 & 0 & 0 & 0 & -1 \end{array}\right]$$

which decomposes in the factors (we choose $D_2 = I$):

$$R_1(\lambda) \Leftrightarrow \left[\begin{array}{cc|cc} \lambda-3 & 0 & 1 & 0 \\ 0 & \lambda-2 & 0 & 1 \\ \hline -1 & 0 & 1 & 2 \\ 0 & -1 & 0 & -1 \end{array}\right].$$

$$R_2(\lambda) \Leftrightarrow \left[\begin{array}{cc|cc} \lambda & d & d & 0 \\ -1/d & \lambda & 0 & 1 \\ \hline d & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

One can check that a factorization with degrees $\delta_1 = 1$, $\delta_2 = 1$, $\delta_3 = 2$, is possible when using a Schur decomposition $\hat{A}_s$ where the order of the eigenvalues 3 and 2 is reversed, but this is left out here. ▽▽▽

The reader must be warned at this point that a real transfer function may be cascaded in complex first degree factors and yet be unfactorable in real degree 2 factors. This has been pointed out in [23] where a counterexample is given.

The computation of the factors $\{A_i, B_i, C_i, D_i\}$ thus involves the following steps (between brackets we give an estimate of the number of multiplications and additions involved in each step):

Step 1: Construction of $\hat{A} = A - BD^{-1}C$ ($n^3 + n^2\delta + \delta^2 n$ operations). This step involves the inversion of D, which is well conditioned when cond. D is close to 1. If cond. $D \gg 1$, then the computed $\hat{A}$ may not be accurate.

Step 2: Computing the Schur decompositions UA_sU^* and $V\hat{A}_sV^*$ ($2\times 15\delta^3$ operations). Backward stable software is available for this step [29].

Step 3: Construction of $Q \triangleq V^*U$ and its LU decomposition (δ^3 and $1/3\delta^3$ operations). Backward stability is ensured when appropriate bounds for acceptability are imposed on cond. Q_l.

Step 4: Construction of T and of the cascaded realization $\{A_c, B_c, C_c, D\} = \{T^{-1}AT, T^{-1}B, CT, D\}$ ($\delta^3 + 2(n+\delta)\delta^2$ operations). Stability is again ensured since cond. $T =$ cond. Q_l.

Step 5: Construction of the factors $\{A_i, B_i, C_i, D_i\}$ from $\{A_c, B_c, C_c, D\}$ ($n^2\delta$ operations). According to Remark 2 we choose $D_1 = D$ and the other $D_i = I$. Then we only have to compute $D^{-1}C_c$, and partition A_c, B_c and $D^{-1}C_c$. Here again cond. D determines the accuracy of the results.

The number of operations correspond to complex computations or real computations depending on the data. In general $\delta > n$ and then clearly the (complex or real) Schur decompositions will demand most of the computing time. When a chosen ordering of poles and zeros does not yield a requested LU decomposition, an updating of the Schur decomposition with another ordering of the eigenvalues will be performed using the previous decompositions. An efficient algorithm to reorder the eigenvalues in A_s and $\hat{A}_s$ [21] requires only about 8δ operations for each permuted pair of eigenvalues. As already mentioned above this may lead to a combinatorial problem.

Remark 3

The possible bad conditioning of D may seriously affect the reliability of the obtained factorization. In such a case one can avoid the inversion of D in Steps 1 and 5 by using generalized eigenvalue routines on the pencil (see [25], [32]):

$$\lambda\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A & -B \\ C & -D \end{bmatrix}.$$

Considering this extension one may say that in general the factorization problem can be solved in a stable way. This of course does not mean that the factors are always well defined. Even though we have control on the errors performed in the computations (i.e., numerical stability), these errors may yield total different answers for the computed factors. This is, e.g., reflected by the sensitivity of the Schur decompositions: small perturbations in A and $\hat{A}$ may cause large deviations in both their eigenvalues and invariant subspaces [19], [20], [28]. This will, e.g., occur when eigenvalues are clustered or repeated. If one is only interested in a synthesis of $R(\lambda)$ then this sensitivity is not important, since the backward numerical stability ensures that the computed factorization corresponds to a transfer function "close" to $R(\lambda)$. $\nabla\nabla\nabla$

V. Connections with the Classical Factorization Theory

In the mathematical literature [6], [11], [15] and in circuit theory [4], [9], [31] special attention has been paid to factorizations of "para J-unitary" and "para J-contractive" transfer functions. Especially the work of Livsic is relevant in this context because he was the first to use and prove the factorizability conditions for this special case. In [4], [9] the relation between factoring these types of matrices and the Darlington synthesis is discussed. Let

$$J \triangleq \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}, \qquad p+q=n$$

be a signature matrix. Let us define the lower star operation by

$$R_*(s) \triangleq [R(-\bar{s})]^*.$$

A matrix $R(s)$ is said to be para J-unitary and para J-contractive if

$$R(s)\cdot J\cdot R_*(s)=J \quad \text{and} \quad R(s)\cdot J\cdot[R(s)]^* \leq J, \qquad \text{for } \operatorname{Re}\{s\}>0.$$

It can be shown, using [11], [15], that any para J-unitary and para J-contractive matrix can be realized as

$$R(s)=D\left[I_n - B^*(sI_\delta - A)^{-1}BJ\right]$$

where

$$D^*JD=J \quad A+A^*=-BJB^*.$$

The above state-space model is also "J-balanced." It is easy to check that the zero operator of this model is given by

$$\hat{A} \triangleq A+BJD^{-1}DB^* = -A^*$$

so that any upper Schur form for A is a lower Schur form for $\hat{A}$, and the matrices U and V in (28) are equal. The subspaces $\mathcal{X}_i$ and $\mathcal{Y}_i$ can, therefore, be chosen orthogonal. The matrix $R(s)$ is thus factorable for any ordering $\{a_{ii}\}$ of poles and corresponding ordering $\{-\bar{a}_{ii}\}$ of zeros. Moreover the condition number of T becomes one and no numerical difficulty is encountered to perform the factorization. For the discrete time case similar results hold, and they can be derived from the above using a bilinear transformation (see [6] for the case $J=I$).

VI. Conclusion

In this paper we have discussed the problem of cascade factorization from a computational point of view. Necessary and sufficient conditions have been derived for the existence of a cascade factorization with factors of arbitrary degree and a numerically sound criterion to check this condition has been given. A relatively simple algorithm to compute the (real or complex) factors has been derived.

The material is based on an approach developed in earlier work [2], [3], [8], [26]. From this theory other related results can be retrieved, such as the classical factorization theory for para J-unitary transfer functions [4], [6], [9], [11], [15], [31] (see Section V), spectral factorization methods [1], [13], [30] (see [3]) and the conditions for factorizability using the parameters of the transfer function [10], [23], [24] (using realizations of the type discussed in [27]).

An important issue in cascade synthesis is the sensitivity of certain properties of the transfer function (poles, zeros, impulse response) with respect to perturbations in the parameters of the factors. Assuming that one starts out with a realization with low sensitivity, then the sensitivity of the cascaded realization will be proportional to the condition number of the required state-space realization T. For this reason one wants to find the ordering of poles and zeros that minimizes cond. T. This also shows the importance of choosing an original state-space model with low sensitivity (e.g., balanced realization [14]). We conclude by mentioning some open problems in factorization theory. It is known that prime factors—i.e., factors which cannot be minimally factored themselves—of degree two and three exist [10], [24]. It is not known whether prime factors of any dimension larger than or equal to two and any degree exist. It is also known that some matrices have two or more minimal factorizations into prime factors for which the degrees of the factors do not match up. This is in contrast to factorizations in a ring or module where prime factors can be different but their degrees will match. It is not known for a given matrix, which and how many factorizations exist.

References

[1] B. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis. A Modern Systems Approach*. NJ: Prentice Hall, 1972.

[2] H. Bart, I. Gohberg, and M. Kaashoek, "*Minimal factorization of matrix and operator functions*," Birkhauser, Basel, 1979.

[3] H. Bart, I. Gohberg, M. Kaashoek, and P. Van Dooren, "Factorizations of transfer functions," *SIAM J. Contr. Opt.*, vol. 18, pp. 675-696, Nov. 1980.

[4] V. Belevitch, *Classical Network Theory*. San Francisco, CA: Holden Day, 1968.

[5] A. Bjorck and G. Golub, "Numerical methods for computing angles between linear subspaces," *Math. Comp.*, vol.. 27, pp. 579-594, July 1973.

[6] M. Brodskii, "Unitary operator colligations and their characteristic functions," *Russ. Math. Surv.*, vol. 33, pp. 159-191, 1978.

[7] C. Chen, *Introduction to Linear System Theory*. New York: Bolt Rinehart & Winston, 1970.

[8] P. Dewilde, "Minimal multivariable cascade synthesis: a structure theory," in *Proc. 20th Midwest Symp. Circuits Syst.*, Western Periodicals, pp. 782-784, Aug. 1977.

[9] P. Dewilde, V. Belevitch, and R. Newcomb, "On the problem of degree reduction of a scattering matrix by factorization," *J. Franklin Inst.*, vol. 291, pp. 387-401, May 1971.

[10] P. Dewilde and J. Vandewalle, "On the factorization of a nonsingular rational matrix," *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 637-645, Aug. 1975.

[11] M. Livsic, "Operators, oscillations, waves (open systems)," *Amer. Math. Soc.*, 1972.

[12] B. McMillan, "Introduction to formal realizability theory I & II," *Bell Syst. Tech. J.*, vol. 31, pp. 217-279, pp. 541-600, 1952.

[13] B. Molinari, "Equivalence relations for the algebraic Riccati equation," *SIAM J. Contr.*, vol. 11, pp. 272-285, May 1973.

[14] B. Moore, "Singular value analysis of linear systems," in *Proc. 17th IEEE Conf. Dec. Contr.*, pp. 66-73, Jan. 1979.

[15] V. Potapov, "The multiplicative structure of J-contractive matrix functions," *Amer. Math. Soc. Transl.*, Ser. 2, vol. 15, pp. 131-244, 1960.

[16] H. Rosenbrock, *State-space and Multivariable Theory*. New York: Wiley, 1970.

[17] L. Sachnovic, "On the factorization of an operator-valued transfer function," *Sov. Math. Dokl.*, vol. 226, pp. 203-207, 1976.

[18] B. Smith *et al.*, *Matrix Eigensystem Routines. EISPACK Guide*. New York: Springer, 1976.

[19] G. Stewart, *Introduction to Matrix Computation*. New York: Academic, 1973.

[20] ——, "Error and perturbation bounds for subspaces associated with certain eigenvalue problems," *SIAM Rev.*, vol. 15, pp. 727-764, Oct. 1973.

[21] ——, "Algorithm 506: HQR3 and EXCHNG. Fortran subroutines for calculating and ordering the eigenvalues of a real upper Hessenberg matrix," *ACM TOMS*, vol. 2, pp. 275-280, Sept. 1976.

[22] A. van der Sluis, "Condition numbers and equilibration of matrices," *Num. Math.*, vol. 14, pp. 14-23, Dec. 1969.

[23] J. Vandewalle and P. Dewilde, "On the irreducible cascade synthesis of a system with a real rational transfer matrix," *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 481-494, Sept. 1977.

[24] ——, "Local I/O structure theory for multivariable systems and applications to cascade realization," *IEEE Circuits Syst.*, vol. CAS-25, pp. 279-289, May 1978.

[25] P. Van Dooren, "The generalized eigenstructure problem. Applications in linear system theory," Ph.D. dissertation, Kath. Univ. Leuven, Louvain, Belgium, May 1979.

[26] P. Van Dooren and P. Dewilde, "Minimal factorization of rational matrices," in *Proc. 17th IEEE Conf. Dec. Contr.*, pp. 170-172, Jan. 1979.

[27] P. Van Dooren, P. Dewilde, and J. Vandewalle, "On the determination of the Smith-McMillan form of a rational matrix from its Laurent expansion," *IEEE Circuits Syst.*, vol. CAS-26, pp. 180-189, Mar. 1979.

[28] J. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford: Clarendon, 1965.

[29] J. Wilkinson and C. Reinsch, "Handbook for automatic computations II: Linear algebra," Berlin: Springer, 1971.

[30] J. Willems, "Least squares stationary optimal control and the algebraic Riccati equation," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 621-634, Dec. 1971.

[31] D. Youla and P. Tissi, "N-port synthesis via reactance extraction. Part I," in *1966 IEEE Int. Conv. Rec.*, pp. 183-208, 1966.

[32] P. Van Dooren, "The generalized eigenstructure problem in linear system theory," *IEEE Trans. Automat. Contr.*, vol. AC-26, Feb. 1981.

Paul M. Van Dooren (S'79–M'80) was born in Tienen, Belgium, on November 5, 1950. He received the engineering degree in computer science and a doctorate in applied sciences, both from the Katholieke Universiteit, Leuven, Belgium, in 1974 and 1979, respectively.

From 1974 to 1977 he was Assistant at the Division of Applied Mathematics and Computer Science of the Katholieke Universiteit of Leuven. He was a Research Associate at the University of Southern California, in 1978-1979 and a Post-Doctoral Scholar at Stanford University, in 1979-1980. He is currently with the Philips Research Laboratory, Brussels, Belgium. His main interests lie in the areas of numerical analysis and linear system theory.

Patrick Dewilde (S'66–M'68) was born in Korbeek-Lo, Belgium, on January 17, 1943. He received the engineer degree from the University of Louvain, Louvain, Belgium, in 1966, and the Ph.D. degree from Stanford University, Stanford, CA, in 1970.

He has held research and teaching positions with the University of Louvain, the University of California, Berkeley, and the University of Lagos, Nigeria. He is currently Professor of Network Theory at the T.H. Delft, The Netherlands. His main interests are in applied algebra (systems and network synthesis, factorization properties for several types of matrix functions, scattering matrix theory, and numerical analysis) and in teaching applied algebra.